

# Measurement based entanglement under conditions of extreme photon loss

Earl T. Campbell<sup>1</sup> and Simon C. Benjamin<sup>1,2,\*</sup>

<sup>1</sup>Department of Materials, University of Oxford, Parks Road, Oxford OX1 3PH, UK.

<sup>2</sup>Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543.

The act of measuring optical emissions from two remote qubits can entangle them. By demanding that a photon from each qubit reaches the detectors, one can ensure that no photon was lost. But the failure rate then rises quadratically with loss probability. In [1] this resulted in 30 successes per billion attempts. We describe a means to exploit the low grade entanglement heralded by the detection of a lone photon: A subsequent perfect operation is quickly achieved by consuming this noisy resource. We require only two qubits per node, and can tolerate both path length variation and loss asymmetry. The impact of photon loss upon the failure rate is then linear; realistic high-loss devices can gain orders of magnitude in performance and thus support QIP.

Recent experimental successes [1, 2, 3, 4] demonstrate that measuring two quantum systems can cause them to become entangled. Consider an optically active matter system: an atom, or an atom-like structure such as a quantum dot or a crystal defect. Following laser stimulation such a system may emit a photon; its internal state is then correlated with the photon. We may generate entanglement by stimulating two separate systems simultaneously and monitoring their emissions in such a way that we determine characteristics of their mutual state *without* learning the source of any given photon [5] (e.g. Fig. 1). This approach, and certain comparable alternatives [8, 9], have the profound advantage that the component systems can be far apart, which may make a *scalable* technology more attainable.

Any photon loss from the apparatus will leave the matter systems in an uncertain state. Unfortunately real systems will always suffer from finite photon capture efficiency, and this can be exasperated by non-ideal physics in emission process. For example nitrogen-vacancy (NV) defects in diamond emit less than three percent of their light cleanly, i.e. without giving rise to phonons. A scalable technology would need further lossy components for optical routing. Photon loss is therefore to be recognised as a fundamental difficulty.

One solution is *weak excitation*: reduce the rate of photon generation so that emission from both qubits simultaneously is very rare. A detector click is taken to imply that exactly one photon was emitted, and successfully reached the detectors [5, 6, 7]. However the rate of entanglement generation is then inversely linked to the fidelity. Entanglement between atomic ensembles has been accomplished in this way [10]. In alternative *two-photon* schemes both matter systems emit a photon, and success is heralded by detecting them both [11, 12, 13, 14, 15, 16]. We register a failure if either photon is lost, so the success rate falls as the square of the transmission probability. Recently two macroscopically separate atoms were entangled in this way [1] at a success rate of about thirty per billion attempts (with a recent preprint reporting a 13 fold improvement [4]).

Here we describe a new approach to entanglement generation that can be fundamentally more rapid. Adopting an en-

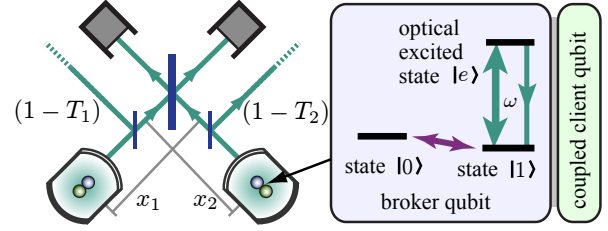


FIG. 1: Schematic of an apparatus suitable for creating entanglement between distant entities, using a beam splitter for erasure of *which path* information. Our model incorporates uncertain optical path length and asymmetric photon loss via the parameters shown. Photon loss is modelled by beam splitters with transmittance  $T_1$  and  $T_2$ ; loss at different parts of the apparatus has an equivalent description in terms of  $T_1$  and  $T_2$ . We define  $\phi$  to be the transmittance asymmetry, such that  $\sin(2\phi) = (T_1 - T_2)/(T_1 + T_2)$ . Asymmetry in path length is parametrised by  $\Delta = \pi(x_1 - x_2)/\lambda$ .

tanglement distillation technique, we develop a protocol that can quickly implement a perfect gate operation using the imperfect entanglement that results from seeing a lone detector ‘click’. Our gate operation is a parity projection, which supports universal quantum computing through the graph state approach; we evaluate the performance in this context. The approach requires two qubits at each local site. This modest level of complexity has already been demonstrated in several systems: In atom trap devices one can confine two atoms, while recent experiments on NV defects have shown the interacting electron and nuclear spins are individually controllable [17]. The utility of such two-qubit nodes has been studied in the context of quantum repeaters [18].

In order to provide a clear exposition, we will assume the specific energy level structure shown in Fig. 1. Alternatives such as  $\Lambda$  structures can be equally suitable. We begin by preparing each of the optically active qubits, which term *brokers* following Ref. [20], in state

$$|\theta\rangle = \cos(\theta)|0\rangle + \sin(\theta)|e\rangle. \quad (1)$$

Here  $|e\rangle$  is the state that decays radiatively to  $|1\rangle$ , see Fig. 1. We collect photons using a lens or a cavity system and direct them through a beam splitter. Given an ideal apparatus, if exactly one photon is detected then the brokers are projected onto the  $(|01\rangle + |10\rangle)/\sqrt{2}$  state (neglecting any phase that de-

\*Electronic address: s.benjamin@qubit.org

depends on *which* detector clicked, which is trivially corrected by a local operation). However in reality we may have non-number resolving detectors, photon loss (which may be asymmetric) and path length variations. Hence, a single click corresponds to a mixed state

$$\rho_B = (1 - \eta) \mathcal{Z}_{B1}^{\phi, \Delta} |\Psi^+\rangle \langle \Psi^+| \mathcal{Z}_{B1}^{\phi, -\Delta} + \eta |11\rangle \langle 11|, \quad (2)$$

where operator  $\mathcal{Z}_{B1}^{\phi, \Delta}$  represents the effects of asymmetry in the apparatus, as parameterized in Fig. 1. This operator can be formally represented as acting on broker qubit  $B1$  alone:

$$\mathcal{Z}_{B1}^{\phi, \Delta} = [\cos(\phi) \mathbb{1} + \sin(\phi) Z_{B1}] [\cos(\Delta) \mathbb{1} + i \sin(\Delta) Z_{B1}]. \quad (3)$$

In Appendix A we show that the probability that we indeed see a click, i.e. that we obtain  $\rho_B$  from initial state  $|\theta\rangle|\theta\rangle$ , is

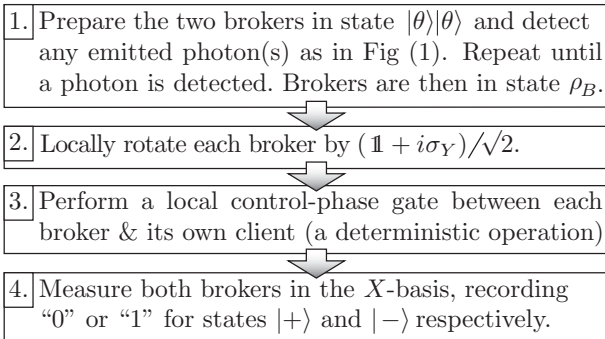
$$P_{\text{click}} = T \sin^2(\theta) (2 - T \sin^2(\theta) \cos^2(2\phi)) \quad (4)$$

where  $T = (T_1 + T_2)/2$ . The weighting in  $\rho_B$  is shown to be

$$\eta = \frac{\sin^2(\theta) [2 - T \cos^2(2\phi)]}{2 - T \sin^2(\theta) \cos^2(2\phi)}. \quad (5)$$

From the work of Bennett *et al.* [22] we know that there is a finite chance of distilling a perfect Bell state from any two states of the form  $(1 - \eta) |\Psi^+\rangle \langle \Psi^+| + \eta |11\rangle \langle 11|$ . Note this is equivalent to (2) without imperfections  $\Delta$  and  $\phi$ . Here we show that one can perform a perfect *parity projection* by consuming two (or more) resource states of the general form (2). We find that  $\Delta$  and  $\phi$  can be completely unknown provided that they do not drift significantly over the course of the protocol, and that two qubits per node suffice for all operations. Thus we obtain the high fidelity typical of two photon schemes, but at a rate that is affected only linearly by  $T$ .

We present our protocol in terms of a basic *iterate* that must be performed more than once. The aim is to perform a single high fidelity parity projection on two remote *client* qubits. Note that although we are adopting the *broker-client* terminology from Ref. [20], that earlier paper did not exploit the brokers for distillation and thus implicitly assumed that the initial entanglement mechanism is near perfect ( $\eta \rightarrow 0$  and  $\mathcal{Z} \rightarrow \mathbb{1}$  in Eqn.(2)). The first flow chart shows one iterate:



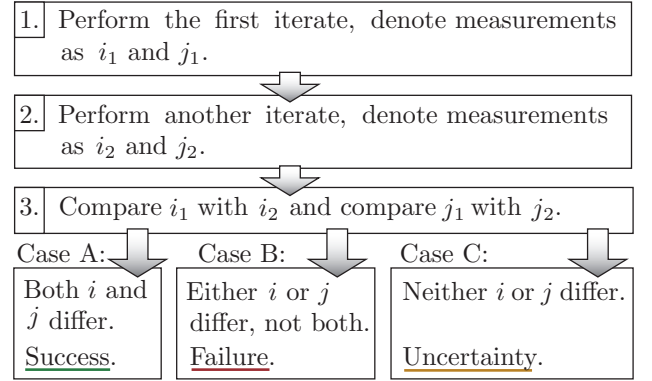
Step 2 will map  $|11\rangle \rightarrow |++\rangle$  and  $\mathcal{Z}_{B1}^{\phi, \Delta} |\Psi^+\rangle \rightarrow \mathcal{X}_{B1}^{-\phi, -\Delta} |\Phi^-\rangle$ , where  $|\Phi^-\rangle = (|00\rangle - |11\rangle)/\sqrt{2}$  and

$$\mathcal{X}_{B1}^{\phi, \Delta} = [\cos(\phi) \mathbb{1} + \sin(\phi) X_{B1}] [\cos(\Delta) \mathbb{1} + i \sin(\Delta) X_{B1}].$$

Using  $i$  and  $j$  to denote measurement results on the brokers, one complete iterate will transform clients qubits from an initial state  $\rho_C$  to an (unnormalised) final state (see Appendix B):

$$E_{i,j}(\rho_C) \equiv \frac{1}{2} (1 - \eta) \mathcal{Z}_{C1}^{\phi_i, \Delta_i} P_{i,j} \rho_C P_{i,j} \mathcal{Z}_{C1}^{\phi_i, -\Delta_i} + \eta S_{i,j} \rho_C S_{i,j}, \quad (6)$$

where  $P_{i,j}$  is a parity projector with parity opposite to that of  $i, j$ , and  $S_{i,j}$  is an undesired projector onto separable states. Specifically  $S_{i,j} = |i, j\rangle \langle i, j|$  while  $P_{i,j} = (\mathbb{1} - (-1)^{i+j} Z_{C1} Z_{C2})/2$ , the labels  $C1$  and  $C2$  referring to the two client qubits. The variables  $\phi_i$  and  $\Delta_i$  are  $-(-1)^i \phi$  and  $-(-1)^i \Delta$ , respectively. Since this state is mixed, we will require least one more iterate. Initially we will consider two iterates in total, moving to more general cases later.



In Case A the we have measurement results of the same parity, but differing specific  $i, j$  values. The corresponding parity projectors are consistent,  $P_{i_n, j_n} P_{i_1, j_1} = P_{i_1, j_1}$ , but all other projector combinations vanish,  $P_{i_n, j_n} S_{i_1, j_1} = S_{i_n, j_n} P_{i_1, j_1} = S_{i_n, j_n} S_{i_1, j_1} = 0$ . We conclude with certainty that the clients qubits have simply been acted on by  $\mathcal{Z}_{C1}^{\phi_i, \Delta_i} \mathcal{Z}_{C1}^{-\phi_i, -\Delta_i} P_{i_1, j_1}$ , which is proportional to a pure parity projection, as  $\mathcal{Z}_{C1}^{\phi_i, \Delta_i} \mathcal{Z}_{C1}^{-\phi_i, -\Delta_i} = \cos(2\phi) \mathbb{1}$ . The probability of reaching Case A after two iterations is

$$P_{\text{two}} = \cos^2(2\phi) (1 - \eta)^2 / 2. \quad (7)$$

In Case B the parity of  $i, j$  has changed and consequently  $P_{i_n, j_n} P_{i_1, j_1} = 0$ , so we must conclude that the client qubits have been projected into a separable state. This is an unrecoverable failure and the clients must be reset before trying again (any prior entanglement with other qubits is lost). In Case C the quantum operation on the clients still contains some asymmetry operator,  $(\mathcal{Z}_{C1}^{\pm \phi_i, \pm \Delta_i})^m$ , and will also produce a mixture if all  $i, j$  have been the same. We could choose to abort here, declaring Case C a failure, which would be a 2 *iterates* only (2IO) strategy. Later we will show that one can also persist with further iterations in order to resolve the uncertainty.

In Fig. 2 we characterise the performance in the simplest scenario: the 2IO protocol acting on clients initialised to  $|+\rangle|+\rangle$ , thus producing Bell pairs at a rate  $R$ . The rate is expressed in units of  $1/\tau$ , where  $\tau$  is the time for a single attempt at generating a photon. Rate  $R$  is optimised by a specific  $\theta$  as shown in Fig. 2(a) and detailed in Appendix D. In

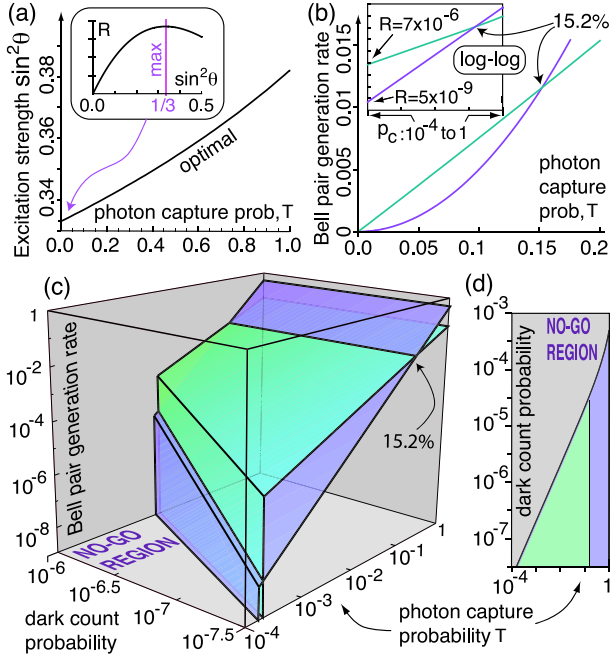


FIG. 2: Attainable rate of Bell pair production,  $R$ . (a) Optimising experimental parameter  $\theta$  (corresponding to amplitude of  $|e\rangle$  in initial state  $|\theta\rangle$ ) to maximise  $R$ . (b) Linear and log-log (inset) plots comparing the rate of Bell pair production in units of  $1/\tau$  using our approach (green) versus a comparable prior scheme, i.e. the double heralding protocol [14] (blue). Other schemes involving detecting two-photons [11, 12, 13, 14, 15] will have similar performance in the high photon loss domain. Note the crossover point of 15.2% above which our approach is unhelpful. (c) Regions in which our approach (green) and double-heralding (blue) have superior rates of generation, given a finite probability of a dark count occurring while we monitor for photons following excitation of the matter systems. Here we assume a minimum acceptable fidelity of  $1 - 10^{-3}$ , thus there is a ‘no-go’ parameter region where both approaches fail. (d) A top-down view, over a wider range of dark counts. For simplicity we assume here that photon loss is symmetric:  $T_1 = T_2 = T$ .

Fig. 2(b) we compare our approach to a *two photon* entanglement scheme, finding three orders of magnitude improvement when  $T$  reaches  $10^{-4}$  (the approximate value that has been achieved experimentally [1]).

Dark counts can be a primary cause of infidelity in entanglement achieved by the path-erasure approach [1]. The approach described here, with its single broker-client pair at each location, is vulnerable to accumulation of dark count noise over successive iterations. However because the number of iterations is so small, the accumulation need not be substantial. In Fig. 2(c) we make the same comparison as in 2(b) but now with finite dark counts and insisting on a minimum fidelity of  $1 - 10^{-3}$ . The thresholds defining the ‘no-go’ region are nearly identical: the reference scheme can tolerate marginally higher dark counts, about 5% greater, and still achieve the target fidelity (note the narrow shelf just visible in the 3D plot). Note that to enter the no-go region one could introduce further ancilla qubits and combine the protocol described here with schemes like Refs. [19, 24, 25, 26].

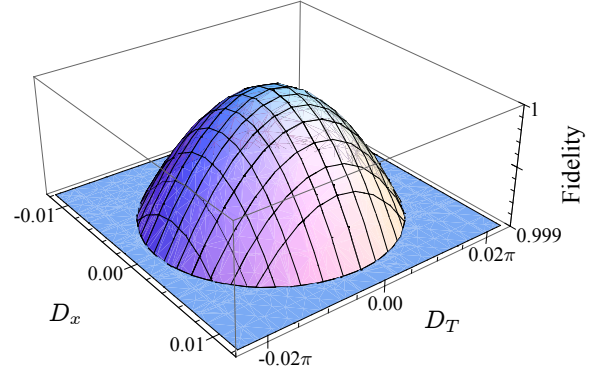


FIG. 3: The fidelity of the 2IO strategy in the presence of drifting apparatus cut off at fidelity  $1 - 10^{-3}$ . Drift in the path length is quantified by  $D_x = ((x_1 - x_2) - (x'_1 - x'_2))/\lambda$ , where unprimed (primed) variables represent the initial (final) values. Drift in the photon transmission probabilities is quantified by  $D_T = 1 - \sqrt{(T'_1 T_2)/(T_1 T'_2)}$ , with the same priming notation.

We have shown that our protocol does not require one to determine any difference in path length between the ‘arms’ of the device, nor any asymmetry in the transmission probabilities; the former will cancel and the latter only reduces the overall probability of success without degrading the entanglement. However, this conclusion is only valid if the unknown quantities do not drift significantly *during* a given instance of the protocol. If indeed such drift has occurred, but is relatively slight, then it is straightforward to characterise the level of resulting error. In Appendix C we perform this calculation and find that the infidelity  $\epsilon = (\pi D_x)^2 + (D_T/2)^2 + \text{Order}(D_x, D_T)^4$ . In Fig. 3 we show the fidelity of the parity projection of a 2IO strategy, where drift occurs between the first and second iterate. Assuming that only path length drift is significant, the infidelity will be below  $10^{-3}$  provided that drift  $D_x < 1/(32\pi)$ .

Our protocol has applications far beyond simply generating high quality Bell pairs. Because we perform a parity projection on the client qubits rather than purifying a specific state, we can use the same protocol to subsequently entangle those qubits with other partners. It is now well understood that one can thus build up a many-qubit resource such as a graph state, and so support efficient universal quantum computation [27, 28, 29]. We will evaluate the performance of our protocol in the context of the specific growth strategy depicted in Fig. 4. There are now two numbers to optimise, the rate of new entanglement generation, and the probability that a given entanglement operation will succeed. Thus it is interesting to generalise our simple 2 *iterates only* protocol in order that ‘Case C’ need not be abandoned. In fact this is straightforward: Over multiple iterates, outright failure occurs only when the parity of measurement outcomes  $i, j$  in the latest iterate differs from the previously seen parity. Success requires that all outcomes have the same parity *and* the two possible instances of that parity have occurred the *same number of times*. The latter criterion assures that any errors described by  $\mathcal{Z}$  will not degrade the resulting entanglement; our

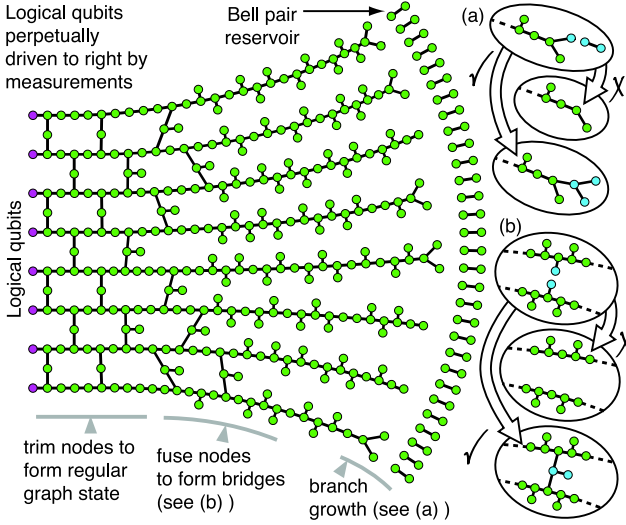


FIG. 4: A suitable method for full scale quantum computation using the well studied *graph state* entanglement. The approach depicted here provides the context for us to evaluate our protocol. Each dot (or ‘node’) represents a qubit, each line (or ‘edge’) represents entanglement between qubits. On the left side the graph state has a regular ordering, using which one can perform a computation purely by measuring out individual qubits (each in an appropriate basis) [27, 28, 29]. The graph is thus consumed from the left, driving the logical qubits to the right such that they are forever on the left-most fringe. Simultaneously a second set of measurements, i.e. those corresponding to the entanglement process described here, generate fresh graph state structure (right edge) and fuse it into a regular array (central). Entanglement is created ‘just in time’ for use. As each physical system is measured out on the left, it is reused in the process of new growth. The small figures to the right show the results of success (creation of new entanglement) or failure (destruction of prior entanglement).

former expression for the state projector simply generalises to  $(\mathcal{Z}_{C1}^{\phi_i, \Delta_i})^{\frac{n}{2}} (\mathcal{Z}_{C1}^{-\phi_i, -\Delta_i})^{\frac{n}{2}} P_{i,j,1}$ , which reduces to a pure parity projection as before. Given a system where  $\mathcal{Z} \rightarrow \mathbb{1}$ , this criterion could of course be relaxed.

In the paradigm illustrated in Fig. 4, one would employ the simple *2IO* strategy for the creation of the Bell pair reservoir, and the more sophisticated multi-iterate protocol for the creation of other ‘edges’ where failure will cause damage (as shown in the Fig. 4 insets). The characteristic time required to perform operations on the logical qubits, i.e. the effective clock speed of the quantum computer, depends on the time needed to create new edges in the graph. It is straightforward to determine this variable in the limit of low photon capture probability  $T$ ; the calculation appears in Appendix E. The average time required to increase a branch length by one qubit, accounting for both successes and failures of the form shown in Fig. 4(a), is found to be  $29.0(\tau/T)$ . This is in contrast to  $\text{Order}(\tau/T^2)$  for two-photon schemes such as Ref. [14]. We emphasise that although this ‘clock speed’ depends on  $T$ , in our approach the probability of success is substantially *decoupled* from  $T$ . Finally we note that although this example is 2D, one can equivalently generate higher dimensional graphs,

including the structures suggested by Raussendorf *et al.* for fault tolerant computation with a high threshold [30].

We have described a form of measurement-based entanglement generation that can be used in situations of high photon loss. Performance can be orders of magnitude higher than current experiments, thus bringing fault tolerance thresholds correspondingly closer and enhancing the prospects for a highly scalable technology. We thank Joe Fitzsimons, Sean Barrett, Jason Smith, Pieter Kok and David Moehring for helpful comments. This research was supported by the Royal Society and the QIP IRC.

#### APPENDIX A: A DERIVATION OF $\eta$ & $P_{click}$ .

We denote  $P_{click}$  as the probability of single detector click from two brokers in the state  $|\theta\rangle|\theta\rangle$ . This is simply the sum of: the probability of a single photon being emitted and captured,

$$P_{(1)} = 2T \sin^2(\theta) \cos^2(\theta); \quad (\text{A1})$$

and the probability of two photons being emitted and *either* being captured

$$P_{(2)} = \sin^4(\theta)[1 - (1 - T)^2]; \quad (\text{A2})$$

Note that the *either* is inclusive since photons bunch and detector are assumed to be non-number resolving. The resulting mixed state has a  $|11\rangle\langle 11|$  contribution with a proportion  $\eta = P_{(2)}/P_{click}$ .

#### APPENDIX B: THE QUANTUM OPERATION $E_{i,j}(\rho_C)$ .

We derive the quantum operation for measurement results  $i$  &  $j$  by considering how each pure state  $|\Psi^+\rangle$  and  $|11\rangle$  affects the clients, and then weighting these outcomes appropriately. Considering  $\mathcal{Z}_{C1}^{\phi, \Delta}|\Psi^+\rangle$  first, the local rotation  $(\mathbb{1} + iY)/\sqrt{2}$  maps this to the state  $\mathcal{X}_{B1}^{-\phi, -\Delta}|\Phi^-\rangle$ . Next a control-Z is performed between broker and client, followed by an  $X$ -basis measurement, resulting in:

$$\langle \pm |_B C Z_B^C = (\langle 0 |_B \otimes \mathbb{1}_C \pm \langle 1 |_B \otimes Z_C) / \sqrt{2}. \quad (\text{B1})$$

It is easy to verify that this operator maps the asymmetry operators between brokers and clients, such that:

$$\langle \pm |_{B1} C Z_{B1}^{C1} \mathcal{X}_{B1}^{-\phi, -\Delta} = \langle \pm |_{B1} C Z_{B1}^{C1} \mathcal{Z}_{C1}^{\mp\phi, \mp\Delta}. \quad (\text{B2})$$

Using  $i$  &  $j$  to represent the measurement outcomes on brokers  $B1$  and  $B2$ , we conclude that the resulting effect on the client qubits is:

$$\begin{aligned} & \langle (-1)^i, (-1)^j |_{B1, B2} C Z_{B1}^{C1} C Z_{B1}^{C1} \mathcal{X}_{C1}^{-\phi, -\Delta} |\Phi^-\rangle (\text{B3}) \\ &= \langle (-1)^i, (-1)^j |_{B1, B2} C Z_{B1}^{C1} C Z_{B1}^{C1} |\Phi^-\rangle \mathcal{Z}_{B1}^{\phi_i, \Delta_i} \\ &= (\mathbb{1}_{C1} \mathbb{1}_{C2} - (-1)^{i+j} Z_{C1} Z_{C2}) \mathcal{Z}_{B1}^{\phi_i, \Delta_i} / 2\sqrt{2} \\ &= P_{i,j} \mathcal{Z}_{C1}^{\phi_i, \Delta_i} / \sqrt{2} \end{aligned}$$



where  $\phi_i = -(-1)^i \phi$  and similarly  $\Delta_i = -(-1)^i \Delta$ . Hence, the  $|\Psi^+\rangle$  component projects the clients into a parity subspace that is the opposite of the measurement outcome, which is a projector that we denote  $P_{i,j}$ .

When we consider the  $|11\rangle$  component, its separability means that it is sufficient to derive the effect of a single broker on its clients, and the same will hold for the second broker. For a single broker the rotation  $(\mathbb{1} + iY)/\sqrt{2}$  maps  $|1\rangle \rightarrow |+\rangle$ , which is followed by a control-Z operation to give:

$$(|0\rangle + |1\rangle Z_C)/\sqrt{2} \quad (\text{B4})$$

so measuring a broker in the state  $|+\rangle(|-\rangle)$  means that its client is projected into the separable state  $|0\rangle(|1\rangle)$ . Taking both brokers in consideration, the measurement signature  $i, j$ , results in the projector  $S_{i,j} = |i, j\rangle\langle i, j|$ .

To construct a quantum operation we begin by considering that  $|\Psi^+\rangle$  and  $|11\rangle$  contributions will occur with probability  $(1 - \eta)$  and  $\eta$ . However, the occurrence of a parity projector is shared between two different measurement signatures, which introduces a factor of  $1/2$ . This gives an (unnormalised) quantum operation of:

$$E_{i,j}(\rho_C) = \frac{1}{2}(1 - \eta)P_{i,j}\rho_C P_{i,j} + \eta S_{i,j}\rho_C S_{i,j}. \quad (\text{B5})$$

Note that the probability of each outcome can be calculated by taking the trace of  $E_{i,j}(\rho_C)$ . As a check it is reassuring to calculate  $\sum_{i,j \in \{0,1\}} \text{tr}[E_{i,j}(\rho_C)] = 1$ .

### APPENDIX C: DRIFT EFFECTS

Since  $\mathcal{Z}^{\phi, \Delta} \mathcal{Z}^{-\phi, -\Delta} \propto \mathbb{1}$ , we know that for two distillation iterates with different parity outcomes asymmetry operators vanish. Of course, this assumes that the asymmetry is constant over the time between two iterates. In this section we consider the errors caused by drift or jitter in the apparatus:

$$\epsilon = 1 - \frac{|\langle \Psi^+ | \Psi_{\text{drift}} \rangle|^2}{\langle \Psi_{\text{drift}} | \Psi_{\text{drift}} \rangle}, \quad (\text{C1})$$

where,

$$|\Psi_{\text{drift}}\rangle = \mathcal{Z}^{\phi + \delta\phi, \Delta + \delta\Delta} \mathcal{Z}^{-\phi, -\Delta} |\Psi^+\rangle \quad (\text{C2})$$

where  $\delta\phi$  and  $\delta\Delta$  quantify the apparatus drift between first and second iterate. This expands out to a function of asymmetry variables  $\phi$  &  $\Delta$ , and their associated drifts  $\delta\phi$  &  $\delta\Delta$ :

$$\epsilon = \frac{\cos^2(2\phi + \delta\phi) \sin^2(\delta\Delta) + \sin^2(\delta\phi) \cos^2(\delta\Delta)}{\cos^2(2\phi + \delta\phi) + \sin^2(\delta\phi)} \quad (\text{C3})$$

This equation does not directly lend itself to an interpretation in terms of physical properties of the apparatus, so we introduce new variables:

$$D_x = ((x_1 - x_2) - (x'_1 - x'_2)) / \lambda \quad (\text{C4})$$

$$D_T = 1 - \sqrt{\frac{T'_1 T_2}{T'_2 T_1}}$$

where unprimed (primed) physical parameters describe the apparatus for the first (second) iterate of distillation. In terms of these variables, we find:

$$\epsilon = \frac{\sin^2(\pi D_x) + \cos^2(\pi D_x) \left(\frac{D_T}{2 + D_T}\right)^2}{1 + \left(\frac{D_T}{2 + D_T}\right)^2}. \quad (\text{C5})$$

Expanding this expression to leading order in  $D_x$  and  $D_T$ , we have a small-error approximation that is quadratic in both variables:

$$\epsilon = (\pi D_x)^2 + (D_T/2)^2 + O[D_x, D_T]^4 \quad (\text{C6})$$

### APPENDIX D: RATE OF EPR PRODUCTION

In this section we detail how to calculate the rate that Bell pairs can be produced. In calculating this rate we have the preparation angle  $\theta$  as a free parameter that we may optimise over. This varies as a function of the overall photon loss, and also the asymmetry in photon loss.

Using the 2IO strategy, the rate of Bell pair production is:

$$R = \frac{1}{2} \cdot P_{\text{two}} \cdot P_{\text{click}} / \tau, \quad (\text{D1})$$

where  $P_{\text{two}}$  is the probability of succeeding after 2 iterations, and  $P_{\text{click}}$  is probability of succeeding at making the noisy Bell pairs required to drive the distillation. The factor of  $1/2$  is present since 2 iterations are required for each attempt. Finally, the variable  $\tau$  is the time taken for each attempt at the optical protocol used to make noisy Bell pairs (i.e. each attempt within step (1) of the iterate). We neglect the time taken to perform the local operations used in the distillation protocol, since we are principally interested in the regime of heavy photon loss where the dominant time cost comes from the many attempts at obtaining  $\rho_B$ . Expressions for the probabilities,  $P_{\text{two}}$  and  $P_{\text{click}}$ , can be found in the main paper, and using these we can obtain the production rate:

$$R = \frac{T \cos^2(2\phi) \sin^2(\theta) \cos^4(\theta)}{2 - T \sin^2(\theta) \cos^2(2\phi)} / \tau. \quad (\text{D2})$$

It is easy to show that in the limit of vanishing  $T \cos^2(2\phi)$ , the optimal of our free parameter is  $\sin^2(\theta) = 1/3$ . For comparison, the strong excitation used in comparable *two photon* schemes would correspond to  $\sin^2(\theta) = 1/2$ . Since we are primarily interested robustness against photon loss, the figures provided in the main paper are optimal assuming that there is no asymmetry  $\phi = 0$ , and that the amount of photon loss is known. If these quantities are not known then the effect on rate will only be slight.

### APPENDIX E: RATE OF CHAIN GROWTH

In this section we show how to reproduce the optimal growth rate when using a loop strategy to add Bell pairs to

a chain. The essential difference between Bell pair growth and chain growth, is that the latter requires consideration of how distillation failures reduce chain length. This consideration means that a loop strategy gives a superior efficiency, as it avoids failure at any cost. The rate at which qubits are added to a chain is:

$$G = \frac{(2P_{loop} - (1 - P_{loop}))P_{click}}{\langle I \rangle \tau}, \quad (E1)$$

where the new variable  $P_{loop}$  and  $\langle I \rangle$ , respectively represent: the probability of success without having to reset the client qubits; and the expected number of iterates until a distillation attempt is concluded by a success or reset. The two terms in the numerator of  $G$  are the success and failure contributions to chain length. Since we are primarily concerned with robustness against photon loss, we will neglect the effects of photon loss asymmetry.

Before calculating the success probability,  $P_{loop}$ , and expected number of iterates until an attempt is concluded,  $\langle I \rangle$ , we must first find the various probabilities for every  $k^{th}$  iterate. Denoting  $k^{th}$  iterate failure and success probabilities by  $P_f(k)$  and  $P_s(k)$ , respectively, it follows that:

$$P_{loop} = \sum_{k=2}^{\infty} P_s(k), \quad (E2)$$

$$\langle I \rangle = \sum_{k=2}^{\infty} k (P_s(k) + P_f(k)).$$

Both success and failure on a  $k^{th}$  iterate, may occur via a number of different sequences of measurement results, which must be summed over, so:

$$P_f(k) = \left[ 2\eta \left( \frac{1-\eta}{2} \right)^{k-1} + \eta(1-\eta)^{k-1} \right] + 2N_f(k)\eta \left( \frac{1-\eta}{2} \right)^{k-1}, \quad (E3)$$

and,

$$P_s(k) = 2N_s(k) \left( \frac{1-\eta}{2} \right)^k. \quad (E4)$$

The failure outcome has two terms representing: the probability of failing after a sequence of prior iterates with identical

measurement results; and the second represents the all other possible combinations of prior measurement results, of which there are  $N_f(k)$ . These terms differ because in the first instance, the density matrix is mixed prior to failure, whereas in the latter case the density matrix is pure prior to failure. As for the probability of success, this again depends on the number of different sequences of measurement results,  $N_s(k)$ , which will obviously only be non-zero for even values of  $k$ .

Lastly, we must find a method of calculating the different valid combinations of measurement results,  $N_s(k)$  and  $N_f(k)$ . We begin by defining a vector  $\mathbf{v}(k)$  which describes number of different  $k-1$  measurement sequences prior to failure. Prior to failure the all measurement results will be of the same parity, but this leaves two possible outcomes. Using the variable  $i$  to denote the difference in the number of the two possible outcomes, the  $(i+1)^{th}$  entry of  $\mathbf{v}(k)$  is the corresponding number of possible sequences with  $i$  different measurements.

Having defined  $\mathbf{v}(k)$ , we calculate it iteratively, by using:

$$\mathbf{v}(k) = M^{k-2} \mathbf{v}(2), \quad (E5)$$

where  $M$  has non-zero elements everywhere except above and below the diagonal, where the elements are 1; Formally,  $M_{i,j} = \delta_{i,j \pm 1}$ . The vector  $\mathbf{v}(2)$ , at which the first success can occur, has only the first entry as unity and the rest are zero. In matrix form, we have:

$$\mathbf{v}(k) = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 1 & 0 & 1 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix}^{k-2} \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \quad (E6)$$

Now  $N_s(k)$  and  $N_f(k)$  follow directly, as the former is simply the first component,  $N_s(k) = \mathbf{v}_1(k)$ , and the failure variable is a sum from  $i=2$  to the penultimate non-zero entry,  $N_f(k) = \sum_{i=2}^{k-1} \mathbf{v}_i(k)$ .

We have outlined all the necessary steps required to calculate  $G$ , and doing this for different values of  $\eta$  will allow use to find the optimal growth rate. Performing this calculation for small photon capture,  $T$ , and negligible capture asymmetry, one finds that the optimal  $G$  is  $G = 0.0345T/\tau$ . The reciprocal of this rate,  $29.0\tau/T$ , gives the average time taken to extend the chain by one qubit.

---

[1] D. L. Moehring et al, Nature **449**, 68 (2007).  
[2] P. Maunz et al, Nature Physics **3**, 538 (2007).  
[3] J. Beugnon et al, Nature **440**, 779 (2006).  
[4] D. N. Matsukevich *et al*, preprint arXiv:0801.2184v1.  
[5] C. Cabrillo et al, Phys. Rev. A **59**, 1025 (1999).  
[6] S. Bose et al, Phys. Rev. Lett **83**, 5158 (1999).  
[7] D. E. Browne, M. B. Plenio, and S. F. Huelga, Phys. Rev. Lett **91**, 067901 (2003).

[8] P. van Loock et al, (2007), arXiv:quant-ph/0701057.  
[9] C. Y. Hu et al, (2007), arXiv:0708.2019.  
[10] C. W. Chou et al, Nature **438**, 828 (2005).  
[11] L. M. Duan and H. J. Kimble, Phys. Rev. Lett. **90**, 253601 (2003).  
[12] X. L. Feng et al, Phys. Rev. Lett **90**, 217902 (2003).  
[13] C. Simon and W. T. Irvine, Phys. Rev. A **91**, 110405 (2003).  
[14] S. D. Barrett and P. Kok, Phys. Rev. A **71**, 060310 (2005).

- [15] Y. L. Lim, A. Beige, and L. C. Kwek, Phys. Rev. Lett **95**, 030505 (2005).
- [16] S. C. Benjamin, J. Eisert, and T. M. Stace, NJP **7**, 194 (2005).
- [17] M. V. G. Dutt et al, Science **316**, 1312 (2007).
- [18] L. Childress *et al*, Phys. Rev. A, **72**, 052330 (2005).
- [19] Jiang *et al*, Phys. Rev. A **76**, 062323 (2007).
- [20] S. C. Benjamin et al, New Journal of Physics **8**, 141 (2006).
- [21] L. M. Duan, B. B. Blinov, D. L. Moehring, and C. Monroe, Quantum Information and Computation **4**, 165 (2004).
- [22] C. H. Bennett et al, Phys. Rev. A **54**, 3824 (1996).
- [23] L. Childress, J. M. Taylor, A. S. Sorensen, and M. D. Lukin, Phys. Rev. Lett. **96**, 070504, 2006.
- [24] W. Dür and H. J. Briegel, Phys. Rev. Lett. **90**, 067901 (2003).
- [25] L.-M. Duan et al, Quant. Info. and Comp. **4**, 165 (2005).
- [26] D. K. L. Oi, S. J. Devitt, and L. C. L. Hollenberg, Physical Review A **74**, 052313 (2006).
- [27] M. V. den Nest, W. Dür, A. Miyake, and H. J. Briegel, New Journal Physics **9**, 204 (2007).
- [28] R. Raussendorf, D. E. Browne, and H. J. Briegel, Phys. Rev. A **68**, 022312 (2003).
- [29] M. Hein et al, (2006), quant-ph/0602096.
- [30] R. Raussendorf, J. Harrington, and K. Goyal (2007), quant-ph/0703143.